Tables: To Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations. Truman Lee Kelley, Austin, Texas. Bulletin of the University of Texas.

In these tables Professor Kelley has supplied a substitute for most of the numerous arithmetic operations which occur in recurring fashion in the computation of partial coefficients of correlation. In computing the coefficients of regression, the indices of higher order are obtained from those of the first order by means of three fundamental equations of the following form:

$$r_{12.34\cdots n}=\frac{r_{12.34\cdots(n-1)}-r_{1n.34\cdots(n-1)}\,r_{2n.34\cdots(n-1)}}{\sqrt{1-r^2_{1n.34\cdots(n-1)}}\,\sqrt{1-r^2_{2n.34\cdots(n-1)}}},\tag{1}$$

$$\sigma_{1.23\cdots n}=\sigma_1\sqrt{1-r^2_{12}}\,\sqrt{1-r^2_{13.2}}\cdots\sqrt{1-r^2_{1n.23\cdots(n-1)}},\tag{2}$$

$$b_{12.34\cdots n}=r_{12.34\cdots n}\frac{\sigma_{1.34\cdots n}}{\sigma_{2.34\cdots n}}.\tag{3}$$

It is to be noted that the formula (1) for obtaining a partial coefficient of correlation of any order from those of the next lower order is of the same form for all orders. To this end Professor Kelley has computed the values of $\frac{1}{\sqrt{1-a^2}\sqrt{1-b^2}}$ and $\frac{ab}{\sqrt{1-a^2}\sqrt{1-b^2}}$ for all values of a and b from 0 to 1 for equal differences of .01—in all, 20,000 tabular entries. Using equation (1), these tables facilitate the computation of successive orders of partial coefficients, starting with total coefficients of correlation.

The author gives another table for values of r, r^2, $1-r^2$, $\sqrt{1-r^2}$, $\frac{1}{\sqrt{1-r^2}}$, $\log\sqrt{1-r^2}$, $\log\frac{1}{\sqrt{1-r^2}}$, for all values of r from 0 to 1 for equal differences of .01. Using equation (2), this facilitates obtaining the standard deviation of any order from the standard deviation of lowest order. From these the regression coefficients are obtained at once by formula (3). The procedure involves getting first r, then σ, then b.

In his preface, the author promises the extension of the tables for smaller subdivisions of the variables, carried out also to further places of decimals. Though such tables would be more serviceable at the extreme values of the variables, it is questionable whether or not they would lead to more substantial results in forming conclusions based upon the coefficients of correlation. A difference of .001, for example, in any coefficient of correlation, is utterly meaningless to anyone.

In his discussion on "The Function of Partial Coefficients or Correlation and Regression Equations," Professor Kelley illustrates the use of his tables by several examples in detail. He presents a very convenient tabular form for the practical computation of regression equations, with a

detailed discussion of approximations to the regression equation for cases involving a large number of variables.

The presentation, however, adopts poor pedagogic procedure. The apparent implicit confidence in the value of these partial coefficients of correlation, though quite orthodox, seems calculated to create in the mind of the student of statistics a blind faith in the subtle power of this form of analysis. This would tend to make the use of coefficients of correlation simply mechanical, and to that extent eliminate the supreme advantage of naive statistical thought. One would fail to realize that the regression equation is not the final word in any statistical analysis. At best, it is but a convenient approximate representation of the relation between a variable and other variables which partially determine it. Though the tables of themselves are valuable, the author's comments would tend to be misleading to one who is unfamiliar with the precise arbitrary nature of a regression equation. Moreover, regressions are sometimes used for purposes of interpolation for such values of the variables where the actual regression equation is anything but linear. This should be guarded against. Such comments would of course lose their force when applied to what we may term regression hyper-surfaces, and the associated correlation ratios.

The author vaguely refers to $\sqrt{1-r^2}$ as the "coefficient of independence," and to $\frac{1}{\sqrt{1-r^2}}$ as the "measure of importance," "measure of significance," etc., of variables, though r is the correlation with but one other variable. In the second part of a regression equation a variable may be said to be "important" *by definition* in proportion to the size of the weights. It is poor pedagogy to use verbal interpretations of loose significance for functions of precise mathematical form. The latter is at best but an arbitrary means for the classification of the data.

Professor Kelley states the theorem that the determination of the regression equation on the principle of least squares is equivalent to that of determining a linear function of the independent variables which will yield values that correlate most highly with the corresponding values of the dependent variable itself. He refers the student to Yule's "Introduction" for an elementary proof. On referring to that book, however, the student will find nothing in the nature of a proof, though he may be satisfied by a casual statement of the author, without proof, to that effect.

The last sentence of the author's remarks is somewhat cryptic. One would be curious to discover the intended connotation.

HARRY LANGMAN.

New York City.